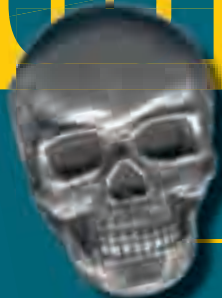
TUTTO QUELLO CHE GLI ALTRI NON DICONO
2€
NO PUBBLICITÀ
SOLO INFORMAZIONE E ARTICOLI
n. 186
www.hackerjournal.it
HACKER JOURNAL
CRACKING
LA TV DIGITALE
CRACCATA
INTERVISTA
LUCI E OMBRE
NELL' NSA
HACKING
OPENWRT
ROUTER EVOLUTI
PROGRAMMING
L'IMPORTANZA
DEL SORT
QUATTORD. ANNO 9 - N° 186 - 8/21 OTTOBRE 2009 - € 2,00
FOCUS ON
VIOLARE LE PASSWORD DI OFFICE
DOCUMENTI TOP SECRET
WLF PUBLISHING

hack·er (hăk’ər)
“Persona che si diverte ad esplorare i dettagli dei sistemi di programmazione e come espandere le loro capacità, a differenza di molti utenti, che preferiscono imparare solamente il minimo necessario.”

Anche noi…

“Le auguro che appena suo figlio avrà accesso a Facebook venga intercettato dai pedofili e che lo incontrino sotto scuola”. (Gabriella Carlucci, rivolgendosi al giornalista Alessandro Gilioli)

Diversamente da quello che pensano molti politici nostrani, non abbiamo mai pensato che Facebook e i social network in generale possano essere un covo di pedofili. Fanno parte di Internet come tante altre cose e sappiamo che la Rete non è altro che lo specchio della società reale: c’è di tutto. Malgrado questa considerazione siamo sempre stati distanti dai social network e da Facebook in particolare perché pensiamo tutt’ora che questi strumenti di comunicazione siano qualcosa di adatto a un pubblico generale e non a un gruppo di tecnici, di hacker, di persone che fanno del pensiero laterale uno stile di vita.

Malgrado questo atteggiamento siamo stati costretti a rivedere le nostre posizioni quando un nostro collaboratore ha trovato su Facebook un gruppo di nostri sostenitori, creato da qualche appassionato. Abbiamo contattato il fondatore per chiedergli chiarimenti ma, senza risponderci mai, ha abbandonato l’amministrazione del gruppo. Così, alla fine, il gruppo l’abbiamo preso noi: se si cerca Hacker Journal su Facebook si può diventare fan della nostra pagina o iscriversi al nostro gruppo.

Questo non esclude, naturalmente, un impegno maggiore sul Web: abbiamo deciso di rinnovarci un po’ e stiamo realizzando una nuova versione del sito, che possa dare un vero supporto ai nostri lettori e che possa essere uno strumento su misura per le nostre esigenze.

Il tutto, ovviamente, senza perdere di vista il fatto che Hacker Journal è principalmente una rivista con la missione esplicita di indagare, provare e informare su tutte quelle cose dell’information technology che gli altri considerano scomode.

Khamul

HACKER JOURNAL: INTASATE LE NOSTRE CASELLE

Diteci cosa ne pensate di HJ, siamo tutti raggiungibili via e-mail, tramite lettera o messo a cavallo... Vogliamo sapere se siete contenti, critici, incazzati o qualunque altra cosa! Appena possiamo rispondiamo a tutti, scrivete!

redazione@hackerjournal.it

Anno 9 – N.186
8/21 ottobre 2009

Editore (sede legale):
WLF Publishing S.r.l.
Socio Unico Medi & Son S.r.l.
via Donatello 71
00196 Roma
Fax 063214606

Realizzazione editoriale
a cura di BMS Srl

Printing:
Roto 2000

Distributore:
M-DIS Distributore SPA
via Cazzaniga 2 - 20132 Milano

Copertina: Daniele Festa

HACKER JOURNAL
Pubblicazione quattordicinale registrata al Tribunale di Milano il 27/10/03 con il numero 601.

Una copia 2,00 euro

Direttore Responsabile:
Teresa Carsaniga

Brunetta connection

In una intervista televisiva, il 21 settembre, Renato Brunetta, ministro della Pubblica Amministrazione, ha fatto una promessa destinata a entrare nella storia dell'informatica.
In positivo o in negativo ma entrerà nella storia dell'IT: da gennaio 2010 le amministrazioni di stato inizieranno a sperimentare la posta elettronica certificata e ai cittadini basterà una mail per evitare code e disguidi. Ovviamente, il ministro ha pensato a chi non ha un computer, la maggioranza degli italiani, promettendo che lo Stato Italiano provvederà a portarglielo e ad occuparsi di chi non lo sa usare.
Intenti certamente lodevoli ma che si scontrano con un mondo, quello della pubblica amministrazione, dove persino l'ottenimento di risposte a raccomandate tradizionali è complicato. Senza contare che, prima di proporre progetti di così ampia portata, sarebbe opportuno fare qualche calcolo. L'ultimo rapporto ISTAT ci dice che il computer è disponibile al 47,8% delle famiglie italiane mentre Internet è diffusa, con varie modalità, nel 37%. Lo stesso ISTAT ci informa che il numero di famiglie italiane era di circa 21 milioni e mezzo nel 2001. Il progetto di cui parla il ministro, quindi, interessa oltre 11 milioni e 220 mila famiglie a cui deve essere dato un computer e oltre 13 milioni e mezzo a cui serve una connessione. In più serviranno corsi indirizzati a utenti che non hanno interesse nell'utilizzo del mezzo. Ci immaginiamo già frotte di pensionati ottantenni che, per non fare la coda all'INPS, si mettono a seguire corsi di informatica generale. Considerando computer di fascia infima, il ministro ha buttato lì un progettino che, se fatto in estrema economia, occuperà oltre 2 miliardi di euro. Esclusa la connessione, naturalmente, perché per l'invio di messaggi PEC risulta fondamentale: chi pagherà l'abbonamento? Chi terrà i corsi? Che corsi saranno e quanto costeranno? Speriamo non siano come la tanto decantata Patente Europea per l'Informatica, capace di creare utenti che non sono in grado di adattarsi a programmi diversi da Office.
L'annuncio parla di una messa in opera del sistema entro il 2010, praticamente domani. Un passo che appare fatto senza un congruo studio di impatto della soluzione adottata: come i sistemisti sanno, i progetti che nascono senza uno studio approfondito non approdano facilmente da qualche parte e spesso falliscono miseramente dopo aver sprecato mucchi di soldi.

UNA FRECCIA AL CUORE, DI GOOGLE BOOK

Tutta l'editoria europea si oppone a Google Books, questo ora che Mountain Views ha raggiunto un accordo con gli editori statunitensi dopo tre anni di class action. Gli editori del vecchio continente chiedono che l'accordo americano sia limitato agli Stati Uniti e ripropone il progetto Arrow come valida alternativa a Google Books. La FEP, che riunisce gli editori di 26 paesi, ha affermato che l'accordo (il Settlement) per chiudere la class action tra Google e le associazioni di autori ed editori americani, che coinvolge anche qualsiasi opera libraria europea disponibile sul mercato Usa, «non costituisce una soluzione per l'Europa, non può essere applicato al territorio dell'Unione Europea e il suo ambito di applicazione non può essere esteso al di fuori degli Stati Uniti».

QUEL COLABRODO CHIAMATO MICROSOFT

A pochi giorni dalle patch che hanno corretto cinque falle critiche presenti nelle varie versioni di Windows, un nuovo problema emerge sotto forma di una vulnerabilità scoperta in Windows Vista e Windows Server 2008. Il bug risiede nell'implementazione del protocollo Smb, la conseguenza più lieve di un attacco è il crash del computer, ma in alcuni casi è possibile eseguire del codice da remoto. Sono immuni da questo problema Windows 2000 e Windows Xp, i quali usano una diversa implementazione di Smb; quella di Vista è stata riscritta apposta per questo sistema e usata anche per Windows 7. La cosa curiosa è che, sebbene Windows 7 Release Candidate sia affetto da questa vulnerabilità, la versione finale non lo è: ciò significa che Microsoft era già a conoscenza del bug, tanto da correggerlo nella Release To Manufacturing di Seven, ma non si è preoccupata di sviluppare una patch per i sistemi operativi precedenti finché l'exploit non è stato reso pubblico. In definitiva solo gli sfortunati e vessati utenti di Vista e Server 2008 sono dunque esposti alla vulnerabilità, per la quale è ora finalmente allo studio una patch.

SILVESTRO MASCHERATO DA TITTI

Twitter, forse più di Facebook, in poco più di tre anni ha conquistato milioni di utenti e migliaia di vip internazionali che lo usano per comunicare direttamente con i propri fan. Purtroppo, proprio per questo, il problema del furto delle identità, è diventata una delle priorità per i gestori del sito che avrebbero deciso di sviluppare una versione più moderna di Twitter, capace di verificare l'identità dell'utente e tutelare i personaggi famosi dalle insidie dei truffatori. Tra i vip vittime di falsi profili ci sarebbero anche Britney Spears, il ministro degli esteri inglese David Miliband, il Dalai Lama e persino la Regina Elisabetta. L'ultimo della serie è stato Tony La Russa, l'allenatore di una squadra di baseball americana, che ha citato in tribunale i responsabili di Twitter dopo aver scoperto un suo falso profilo sul sito di microblogging in cui comparivano commenti ingiuriosi sulla morte di due suoi giocatori. Il social network ha precisato che episodi del genere nel futuro non capiteranno più: «Stiamo lavorando per stabilire l'autenticità degli utenti» ha dichiara un anonimo portavoce di Twitter. «Cominceremo con gli account dei personaggi famosi con i quali abbiamo avuto diversi problemi».

HOT NEWS

ISS STORY, NUOVO CAPITOLO

La storia della falla in Internet Information Services si allunga di un altro capitolo, un capitolo che non piacerà ai suoi utenti. Dopo che Microsoft ha confermato la presenza della vulnerabilità relativa al servizio Ftp, pur circoscrivendola a IIS 5 su Windows 2000, in Rete è apparso un exploit in grado di sfruttarla. Il fatto che si potesse trarne vantaggio solo in particolari circostanze sembrava poter ridurre la portata degli attacchi; invece si è scoperto che non è necessario che un utente anonimo abbia i diritti di scrittura perché sia possibile eseguire codice da remoto sul server, né i problemi sono circoscritti a IIS 5: i server con IIS 6, se colpiti, vanno in crash. Dato che per il momento ancora non esiste una patch, per evitare di rimanere vittime di un attacco è consigliabile disabilitare il server Ftp quando non è necessario oppure negare l'accesso agli utenti non autenticati.

CACCIA ALLA VOLPE

Mozilla ha reso disponibili degli aggiornamenti per entrambe le versioni supportate del proprio browser, rilasciando Firefox 3.5.3 e Firefox 3.0.14. Gli update - che includono anche un sistema per segnalare la disponibilità degli aggiornamenti di Flash - risolvono tre vulnerabilità critiche, una moderata e una a bassa priorità. I rischi per gli utenti che non aggiornano vanno dalla possibilità di crash con corruzione della memoria fino all'esecuzione di codice da remoto. Nonostante spinga con ogni mezzo per l'adozione di Firefox 3.5, Mozilla continuerà a supportare la serie 3.0.x ancora per qualche mese, cercando di convincere i più restii a passare all'ultima release.

LINUX L'UNTORE

Il ricercatore Denis Sinegubko, che si occupa di sicurezza, si è imbattuto in un gruppo di circa 100 server Linux infetti. Oltre ai servizi leciti offerti sulla porta 80 da Apache questi server inviano anche del traffico decisamente meno innocuo sulla porta 8080, gestito da un altro web server, nginx. Questi aggiunge alle pagine web legittime degli iframe che diffondono il malware sfruttando due servizi di Dns dinamico che offrono nomi a dominio gratuiti, i quali puntano verso gli indirizzi IP dei server infetti. Quello che non è ancora chiaro è come i server siano stati infettati. Secondo Sinegubko la responsabilità sarebbe da attribuire agli amministratori, non abbastanza accorti da non farsi sniffare la password di root. DynDNS e No-IP, i due servizi di Dns dinamico usati dai server infetti, hanno già provveduto a disattivare i nomi a dominio indicati dal ricercatore russo, il quale tuttavia segnala che ogni ora vengono registrati due nuovi indirizzi proprio allo scopo di diffondere malware tramite web server compromessi.

Lo spione americano...

Si ritorna a parlare di SWIFT (Society for Worldwide Interbank Financial Telecommunications), il colosso che custodisce informazioni sulle transazioni elettroniche tra 7.800 diverse istituzioni finanziarie del mondo. Finito nel mirino delle autorità europee a tutela della della privacy e accusato di spifferare al controspionaggio statunitense dati sensibili sui conti bancari italiani ed europei all'insaputa dei proprietari, il network con base in Belgio è ora oggetto di intermediazione tra USA e UE nel tentativo di giungere a un accordo per la condivisione delle informazioni ivi contenute. Al momento il presidente di turno dell'Unione è impegnato in negoziati tesi a stipulare un accordo tra le due sponde dell'Atlantico, accordo grazie al quale USA e UE potrebbero condividere le informazioni sulle operazioni finanziarie transatlantiche in attesa che il database americano di SWIFT venga trasferito nella sua nuova sede in Olanda.

... e noi paghiamo!

Escono i due nuovi sistemi operativi per le case di Cupertino e Redmont, ma quanto c'è davvero di nuovo???

Ecco le grandi notizie di questo periodo: Microsoft rilascia Vista Service Pack 3 e Apple rilascia Leopard Service Pack 1! Come, non lo sapevate??? O forse avete davvero creduto che le due aziende avessero davvero messo sul mercato due nuovi sistemi operativi?!?! Illusi!!!
Parliamoci chiaro, il nuovissimo 7 non è niente altro che Vista con, finalmente, un buon debug e una buona soluzione di marketing. Tutti noi abbiamo odiato Vista come la morte e ci abbiamo tirato moccoli e lacrime per tutti i problemi che ha dimostrato di avere, quelli di Redmond lo sanno e quindi hanno pensato bene di fare piazza pulita. Hanno preso il sistema operativo più odiato di sempre. l'hanno pulito e rimesso sul mercato con un nome nuovo e, ovviamente, a pagamento. Così facendo hanno fatto felici tutti, l'utente ha finalmente un PC senza l'odiato e bistrattato Vista e i loro uffici contabilità vedono fluire un bel po' di grana che con un SP3 non avrebbero certo visto.
Discorso diverso, ma non troppo, per Apple: Leopard non aveva tutti i problemi di Vista ma era ormai vecchio di un anno, così Steve Jobs e soci pensano bene di mettere sul mercato, a pagamento, una bella nuova versione, ironicamente la chiamano anche in modo simile, Snow Leopard, e tutti gli Applemaniacs corrono a spendere i loro sudati 30 euro circa per quello che poi poteva essere un aggiornamento di sistema gratuito.
Ora, io capisco che il periodo è quello che è, capisco che anche grandi aziende come Microsoft e Apple abbiano bisogno di cash, capisco che ricerca e sviluppo si pagano cari e salati ma quello che mi sfugge è il motivo per cui dobbiamo sempre essere noi a pagare i loro conti. Non vedo perché il costo dei loro errori deve essere sempre e per forza a carico delle nostre tasche. Già ci siamo comprati, per forza, dei computer con su Vista, abbiamo pagato per fare un downgrade a Xp, che avevamo già pagato, e ora paghiamo per andare a 7 che non è nient'altro che Vista rivisitato.
Non so voi ma io passo a Ubuntu!!!

BigG

Fa parte del gioco

Il progresso non avviene a salti ma con l'evoluzione

È normale che ci sia qualche polemica sulla decisione di Apple di rilasciare Snow Leopard a pagamento, così come è normale che i sostenitori della mela trovino insopportabile la cosa.
Chi è da anni un utente Windows è decisamente più abituato agli aggiornamenti a pagamento: da Windows 95 a Windows 98, da 98 a XP, da XP a Vista... Fino a Seven. Da questo punto di vista è bruciante che le pratiche di Microsoft, derise dagli utenti Apple, abbiano fatto scuola.
Il problema, tuttavia, non riguarda gli aggiornamenti a pagamento o meno: riguarda la definizione stessa degli aggiornamenti. Un conto è il rilascio di una patch che chiude un buco nel SO, mentre ben diversa è l'aggiunta di funzionalità, il loro miglioramento, l'adeguamento dell'interfaccia ai suggerimenti che gli stessi utenti forniscono ai produttori. Entrambi i processi sono costosi per le aziende ma nessuno potrebbe far pagare la correzione di problemi di funzionamento o sicurezza. Quello sarebbe sleale oltre che illegale. Il secondo passo, invece, deve obbligatoriamente essere a pagamento. Se la pensassimo diversamente e considerassimo ogni sistema operativo come un'evoluzione dei suoi predecessori, cosa che è per forza, allora le aziende dovrebbero investire in ricerca e programmazione senza avere alcun corrispettivo in cambio. Il risultato sarebbe l'assoluta mancanza di evoluzione dei sistemi operativi e ci ritroveremmo ancora con un prompt del DOS. Allo stesso tempo, nessuno obbliga gli sviluppatori a creare software sulle piattaforme più recenti ma a tutti gli attori in gioco è ben chiaro che restare con lo sguardo al passato conviene solo agli utenti. Senza soldi non si fa innovazione e senza innovazione sarebbe impossibile avere un'evoluzione. Chiariamoci: non avremmo mai avuto un GTA IV se Rockstar Games avesse dovuto fornirlo gratis a tutti gli acquirenti del GTA originale.
Alla fine è vero: noi paghiamo. Ma paghiamo prodotti che non sono comunque confrontabili con quelli che avevamo. Se non ci piace questo gioco... Abbiamo le alternative per evitarlo. Ubuntu e il software Open mi sembrano una risposta adeguata...

Khamul

Tecnologici, sì. Ma dove?

Un'ottima confezione ma, sotto sotto, non sembra funzionare davvero

Ho deciso di interessarmi un po' del mio abbonamento telefonico. Da alcuni anni sono un abbonato Vodafone e mi sono un po' stancato di andare in posta a pagare la bolletta. Così ho chiesto l'addebito sul conto telefonico. Più di un anno fa. La banca dice che aspetta comunicazioni da Vodafone, Vodafone le aspetta dalla banca e io continuo a ricevere il bollettino postale. Poi il colpo di genio: chiedo l'addebito su carta di credito. Però non basta telefonare al call center e dare il numero della carta, così come non basta andare sul sito e seguire qualche procedura online. Devo andare presso un negozio convenzionato e strisciare la carta. Cerca il negozio, vacci di persona, perdi quelle 2 ore… Ma, alla fine, sono felice: ho persino deciso di investire qualche soldo per un'offerta che mi dà fino a 2 Gb di traffico ogni mese, facilmente controllabile dal sito Web.

:: Illuso!

Invece non funziona proprio così: sul sito Web, la situazione del traffico dati effettuato è tutt'altro che aggiornata. Addirittura, il sito computa il traffico realizzato usando una tariffa piena e facendomi rischiare l'infarto: 1 Gb di traffico dati senza alcuna promozione è una cifra che supera di diversi ordini di grandezza i 10 euro iniziali. Dopo qualche tempo ecco arrivare la prima bolletta con addebito sulla carta… Con bollettino allegato. 2 ore perse per ritrovarsi alla situazione di partenza. Chiamo il call center e mi viene spiegato che le modifiche di addebito ci saranno, forse, dalla prossima bolletta mentre per i dati la situazione è strana ma non ci possono far nulla.

A questo punto è legittimo sospettare che ci sia del marcio in Danimarca: possibile che una società ad alta tecnologia come dovrebbe essere un operatore telefonico, che pubblicizza in continuazione mirabili prodotti, che spinge ogni minima stupidaggine tecnologica, abbia un sistema IT che fa acqua da tutte le parti?

:: Forse funziona o forse è esploso…

Purtroppo sì e Vodafone non è che una delle tante aziende che vanta un'organizzazione che nel 99% dei casi funziona perfettamente ma nel restante 1% (si spera) fa acqua in abbondanza. Qualche anno fa ho cambiato casa ed ho chiesto al mio nuovo comune la residenza. Il comune ha avvisato la motorizzazione, per il cambio di residenza anche della patente e dell'auto. A nulla

Io, l'ecopass vorrei pagarlo. Ma il sito non funziona con Firefox e non vi devo nulla?

sono valse telefonate al call center, fax, e-mail e tempo perso recandosi agli sportelli di persona: sono passati 4 anni e non mi è arrivato alcun aggiornamento.

Poi ho scoperto che per la Regione Lombardia io abito correttamente nella mia via ma di un'altra città. Inoltre, la mia macchina risulta una Euro 4 dotata di filtro antiparticolato: una cosa utile per circolare nel centro di Milano senza pagare l'ecopass... Solo che la mia è un'auto diesel Euro 4 senza FAP. Lasciamo perdere, poi, il canone RAI: l'ho sempre pagato e ho provveduto ad avvisare del mio cambio di residenza. Col risultato che, per 2 anni, sono stato "gentilmente sollecitato" a pagarlo perché non credevano che io fossi la stessa persona, con lo stesso codice fiscale che lo pagava, prima, nella mia vecchia residenza. Alla fine l'hanno capita quando ho accennato a denunce per minacce e interruzione di pubblico servizio.

Faccende strane se si considera che pago le tasse nella mia residenza, che il libretto di circolazione dell'auto raccoglie dati che mi hanno dato loro, che dovrebbe esserci un archivio informatico nella pubblica amministrazione con un record che mi riguarda, che non ho fatto nessun mistero del mio trasloco.

:: Dati incrociati

Per quanto mi riguarda ho deciso di disinteressarmi della cosa: non ha senso che io perda un giorno di lavoro per convincere il comune di Milano che devo pagargli l'ecopass. Così come aspetto che la Regione Lombardia mi chieda di pagargli qualcosa per informarli che non sono io quello che stanno cercando. Che si rivolgano al mio omonimo, col mio stesso codice fiscale e la macchina targata uguale che abita 30 Km più in là. A me viene il dubbio di essere un apolide che non sa di esserlo. Sono, tuttavia, in ottima compagnia: io sarò un tipo sfortunato ma è bastato un rapido giro di consultazioni tra gli amici e sembra proprio che problemi di questo genere siano decisamente diffusi. Forse sarebbe il caso che queste grandi società e questi enti dalle mille idee iniziassero a mettersi loro un po' in ordine, prima di chiederci continuamente di adeguarci. La questione sembra mettere allo scoperto un problema di livello generale perché non si salva proprio nessuno. Possibile che queste pachidermiche entità non sappiano trovare il modo di tenersi aggiornati i loro archivi? Possibile che non sappiano trovare procedure adatte per sapere se, per esempio, stanno mandando comunicazioni a persone reali? Basterebbe un colpo di telefono in comune per sapere se una persona abita veramente in un dato posto: la residenza la

Il canone RAI è a posto. Adesso sanno che io sono io ed ho solo cambiato casa. Son bastati 2 anni e giusto qualche telefonata...

Attiva di qua, controlla di là... Non funziona come mi avete detto voi!

gestiscono loro. Possibile che non siano ammessi cambiamenti per via telematica di queste informazioni e che a rimetterci siano sempre gli utenti finali? A conti fatti, l'aspetto di questi colossi è certamente quello dell'evoluzione, dei servizi, della novità, della tecnologia. Slogan: miglioriamo per migliorare e semplificarti la vita ad ogni costo. Nella realtà, le cose non stanno affatto così e il tutto si riduce alla risoluzione di alcuni problemi senza quel controllo del dato che dovrebbe essere alla base di qualsiasi archivio. Una sorta di trasformazione della certezza della Scienza dell'Informazione in una Scienza della dell'Approssimazione con, in più, una buona dose di burocrazia. Ci si preoccupa spesso della privacy, delle norme di tutela dei dati personali, dell'impedire che si possa fare data mining sulle nostre informazioni ma la mia esperienza è nettamente diversa. Non sanno dove abito malgrado gliel'abbia detto non so quante volte. Non sanno com'è la mia auto malgrado mi abbiano dato loro un foglio con le sue caratteristiche. Non vogliono i miei soldi malgrado gli abbia dato più volte i dati necessari per prenderseli da soli. A questo punto mi chiedo a cosa gli serva investire sull'IT se non sanno cosa farsene e riescono a tradurla in qualcosa di peggio dei vecchi archivi cartacei. Forse per evitare la polvere. Ma nient'altro.

Caro dato, ti ordino

Gli algoritmi di ordinamento sono fondamentali nella gestione delle informazioni e conviene quindi impararli

Benché siano uno dei punti più interessanti nell'apprendimento dell'informatica e della programmazione, gli algoritmi di ordinamento vedono la luce in un periodo antecedente a quella che si può chiamare l'era dei computer: matematici di tutto il mondo hanno speso energie per la soluzione di problemi che vedono protagonista l'ordine dei dati. Tra l'altro, lo studio dal punto di vista della programmazione degli algoritmi di ordinamento ha come effetto collaterale quello di toccare e quindi insegnare altri algoritmi fondamentali, pertanto lo studio di queste funzioni è doppiamente importante.

:: Classificazione

Gli algoritmi di ordinamento sono numerosi e ognuno ha le proprie caratteristiche, in termini di complessità, efficienza e tipologia di applicazione. Questo ha portato a una classificazione degli stessi in base a diversi parametri, che devono essere presi in esame al momento della scelta di quello più adatto in una particolare situazione. Innanzitutto, che cosa si può ordinare? Virtualmente tutto, dal semplice dato numerico all'array multidimensionale, al record più complesso ed eterogeneo di un database. L'importante naturalmente è che i dati che devono essere posti in ordine siano un numero diverso da 1 e siano disponibili e conosciuti. La prima classificazione grossolana che si può fare dei diversi algoritmi, in ambito informatico, li pone in due macrogruppi fondamentali: gli algoritmi che operano in memoria e gli algoritmi che operano su un supporto diverso. Molto dipende dalle caratteristiche del sistema su cui girano, naturalmente, ma si può presumere che un algoritmo in grado di immagazzinare in memoria tutti i dati da ordinare e lavorare quindi in RAM sia molto più veloce di un altro algoritmo che invece deve lavorare su disco indicizzando i record dei dati ed elaborando un frammento alla volta.

Un fattore che influenza sostanzialmente la velocità di esecuzione di un algoritmo, oltre al supporto in cui sono immagazzinati i dati e a parità del numero degli stessi, è la sua efficienza. Per efficienza si intende naturalmente il tempo impiegato dall'algoritmo per ordinare un numero n di dati, ma questo è direttamente dipendente da quante operazioni è costretto a compiere per portare a termine il proprio compito.

È provato che la massima efficienza di un algoritmo si ha quando sono necessari O(n log n) confronti per ordinare un insieme n di dati (O sta per "order of computation), mentre il peggiore dei casi si ha quando il numero di operazioni corrisponde a $O(n^2)$: una classificazione definitiva non è possibile in quanto molto dipende dalla natura e dal numero dei dati, pertanto per ogni singolo algoritmo si tendono a definire aree di funzionamento "buone", "medie" e "peggiori". Più l'efficienza di un algoritmo rimane nelle aree "buona" e "media", migliore è l'algoritmo stesso. Come però vedremo, gli algoritmi che hanno risultati migliori sono quelli che adottano metodologie miste e/o adattive, pertanto potremo avere una routine che ha normalmente pessimi risultati ma, se integrata correttamente con un'altra, può ottimizzare i risultati di quest'ultima e, nel complesso, a migliorare notevolmente le prestazioni generali del nostro programma.

Codice 2: Quick Sort
(fonte: Wikipedia)

```
function quicksort(array)
    var list less, greater
    if length(array) ≤ 1
        return array
    select and remove a pivot value
pivot from array
    for each x in array
        if x ≤ pivot then append x to less
        else append x to greater
    return concatenate(quicksort(less),
pivot, quicksort(greater))
```

Altre considerazioni che vengono esaminate nella classificazione degli algoritmi sono la stabilità, cioè la capacità dell'algoritmo di mantenere un'eventuale ordine preesistente dei dati anche dopo aver fornito il risultato (pensiamo per esempio a un'insieme di schede ordinate per indirizzo che rimangono tali anche dopo aver compiuto un secondo passaggio per ordinarle anche per nome) e il metodo usato per l'ordinamento stesso, cioè come i dati vengono confrontati tra loro per stabilirne l'ordine.

:: Il peggiore

Tra gli algoritmi di ordinamento che presentano la struttura più semplice sicuramente bisogna ricordare il Bubble Sort. La sua routine fondamentale scorre l'array dei dati da ordinare (o il file contenente i record), li confronta a coppie e li scambia tra loro se si trovano nell'ordine sbagliato. La sua efficienza, però, è tra le peggiori, specialmente quando il numero n di dati da ordinare è molto alto, in quanto necessita di un numero di scambi sempre pari a $O(n^2)$. Trova quindi poche applicazioni pratiche, ma a livello didattico è insostituibile. Un margine di miglioramento si ha ottimizzando il numero di scambi necessari per ogni passata: al termine della prima, infatti, il numero più grande o più piccolo (in base all'ordine desiderato) si troverà in fondo, già nella sua posizione finale, quindi la passata successiva necessita di n-1 scambi, e così via fino all'ultima passata, che richiede un solo eventuale scambio. In Codice 1 troviamo un listato in pseudocodice che implementa la versione ottimizzata del Bubble Sort.

Ma c'è un algoritmo che riesce anche a fare di peggio: si chiama, non a caso, Stupid Sort, o anche Monkey Sort, e si basa sulla storiella che narra che anche una scimmia, battendo a caso i tasti di una macchina da scrivere in un numero illimitato di tentativi, azzeccherà prima o poi quell'unica combinazione che le permette di riscrivere qualunque opera letteraria esistente. L'algoritmo funziona così: avendo a disposizione tempo e numero di tentativi infiniti, si dispongono a caso i dati da ordinare e se ne verifica l'ordine.

Codice 1: Bubble Sort
(fonte: Wikipedia)

```
procedure bubbleSort( A : list of
sortable items ) defined as:
  n := length( A )
  do
    swapped := false
    for each i in 0 to n - 1  inclusive do:
      if A[ i ] > A[ i + 1 ] then
        swap( A[ i ], A[ i + 1 ] )
        swapped := true
      end if
    end for
    n := n - 1
  while swapped
end procedure
```

Se sono ordinati il lavoro è terminato, altrimenti si ricomincia creando un'altra sequenza casuale, fino a quando non si trova quella giusta.

:: E il migliore

Un algoritmo tra i migliori, molto veloce e spesso di efficienza ottimale, è il Quick Sort. Si tratta di un algoritmo ricorsivo (cioè la sua funzione principale richiama se stessa durante il funzionamento per lavorare su un sottoinsieme dei dati) che comporta anche poco spreco di risorse del computer e per questo è spesso adottato in situazioni pratiche. Funziona così: individuata una soglia di valore, si spostano da un lato di essa tutti gli elementi maggiori e dall'altro quelli minori. Dopodiché, la funzione chiama se stessa due volte per lavorare sui due sottoinsiemi creati, e così via. Ne troviamo una traccia in pseudocodice in Codice 2: non è però un algoritmo stabile (cioè rimuove ogni altro ordinamento eventualmente già presente) e, nelle peggiori condizioni, ha una efficienza che si avvicina a $O(n^2)$, ma si tratta di casi particolari. Inoltre, se la routine presenta un difetto di programmazione, esso ne degrada esponenzialmente le prestazioni data la sua natura ricorsiva. Si tratta comunque di uno

tra i migliori algoritmi di ordinamento. Il Merge Sort è un altro algoritmo di ordinamento dalle buone prestazioni. Ha una routine ricorsiva che suddivide i dati in gruppi fino ad arrivare a coppie, i cui elementi vengono quindi ordinati. Si procede poi a unire tra loro due coppie alla volta, per ottenere un numero inferiore di insiemi, e si ripassa attraverso la funzione, fino a quando tutto l'insieme non è riunito e ordinato. Questo particolare algoritmo può essere affrontato anche all'inverso, cioè considerando l'insieme dei dati come un numero n di microsequenze e unendole tra loro una volta ordinate. Il vantaggio principale di questo algoritmo consiste nel fatto che la sua efficienza è sempre pari a O(n log n).
Un algoritmo molto interessante è denominato Insertion Sort. Si tratta di un algoritmo che funziona meglio quando i dati sono già parzialmente ordinati, quindi di per sé non è particolarmente efficiente quando il numero di dati è ampio; tuttavia proprio questa sua particolarità fa in modo che venga spesso unito a un altro metodo di ordinamento per raggiungere risultati migliori di quelli che si otterrebbero con uno solo dei due. Per esempio, lo Shell Sort adotta un miglioramento della tecnica del Bubble Sort mediante il quale i dati vengono spostati all'interno dell'array a passi più grandi di una posizione e, quando sono abbastanza ordinati, passa a un Insertion Sort tradizionale, che a quel punto può dare il meglio di sé. In Codice 3 troviamo uno schema in pseudocodice dell'Insertion Sort.

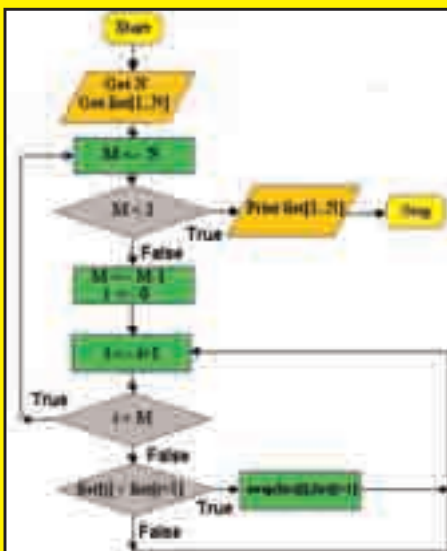
:: Quale scegliere?

Non è facile dare una risposta adeguata a questa domanda: le casistiche sono infinite, ogni situazione presenta esigenze diverse e quindi solamente la sperimentazione può esserci di aiuto. Fatta eccezione per il Monkey Sort e pochi altri, che sono stati ideati più a scopo didattico che altro, tutti gli altri algoritmi hanno pro e contro che variano di molto in base alla natura e al numero dei dati da ordinare. Esistono strumenti software in grado di misurare l'efficienza di un algoritmo in termini di uso della memoria e velocità di esecuzione, sotto forma di routines di profiling che possono essere implementate in un programma per migliorarlo, ma in certi casi sussistono difficoltà addirittura nel compiere queste misurazioni. Per questo motivo non esiste una vera e propria classifica delle prestazioni, ma solamente un'indicazione di massima (quella che abbiamo usato in tutto l'articolo e cosiddetta Big O notation) dell'efficienza di ogni algoritmo. Il medesimo che in una determinata situazione offre il massimo delle prestazioni, cambiando il tipo o il numero di dati può dare risultati pessimi e diventare del tutto inutilizzabile. Tuttavia, un poco di sperimentazione e una buona conoscenza dei principali algoritmi, per lo meno dei più usati in ambito pratico, è indubbiamente molto utile e non devono mancare nel carnet di ogni programmatore, per professione o per diletto. È buona consuetudine quindi scrivere le routine dedicate all'ordinamento dei dati in maniera che si possano velocemente sostituire con altre che adottano un differente algoritmo, per esempio per mezzo di librerie esterne come le DLL di Windows. Se con il primo il programma non dà i risultati sperati, si può sempre tentare con un secondo e valutare poi le differenze tra i due per scegliere il migliore.

Codice 3: Insertion Sort
(fonte: Wikipedia)

```
insertionSort(array A)
begin
   for i := 1 to length(A) - 1 do
   begin
      value := A[i];
      j := i - 1;
      while j >= 0 and A[j] > value do
      begin
         A[j + 1] := A[j];
         j := j - 1;
      end;
      A[j + 1] := value;
   end;
end;
```

A Scuola di programmazione

Lo studio degli algoritmi di ordinamento, come già detto, è molto interessante e avvicina chi lo affronta a diverse tecniche che tornano utili in molti aspetti della programmazione dei computer. Si tratta di concetti che, una volta capiti, possono essere applicati a molteplici algoritmi, o usati per crearne di propri in situazioni specifiche e particolari in cui bisogna risolvere problemi che ancora non sono stati affrontati e risolti da altri. Per esempio, la tecnica di funzionamento del Quick Sort si basa sull'assunto detto Dividi et Impera: è uno dei metodi più usati e consiste nel suddividere il problema affrontato in problemi minori e risolvere prima questi, per poi riunire il risultato nella soluzione di quello più grande. Studiando e applicando in un nostro programma un algoritmo di ordinamento, avremo modo di imparare a maneggiare i dati e la loro struttura, a gestire in maniera ottimale le risorse a nostra disposizione come memoria e potenza di calcolo, ma anche il tempo, che in informatica è sempre un bene preziosissimo.

Samba non è Windows

La diffusione di dischi di rete casalinghi porta a complicazioni inattese

Costano qualche centinaio di euro, promettono di essere la soluzione ai problemi di spazio e si vendono anche nei supermercati: i dischi di rete casalinghi sembrano un miracolo ma possono creare qualche problema. Diversamente dai dischi USB, infatti, i dischi di rete devono avere a bordo qualcosa che gestisca le connessioni e nella stragrande maggioranza dei casi si tratta di ambienti Linux embedded. Motivo per cui, aggiungendo poco hardware, possono trasformarsi in media center calibrati per l'uso casalingo. Nella realtà, tuttavia, l'introduzione di queste tecnologie non è priva di rischi per l'utente. Di solito, infatti, l'interfaccia di configurazione è basata sul Web e l'utente medio non ha la minima idea di come funzionano. Addirittura vengono spesso confusi con normali dischi USB, pur essendo sostanzialmente diversi. Questa confusione presso il pubblico a cui sono destinati e la mancanza di trasparenza verso il loro sistema operativo porta a situazioni in cui i dati risultano a rischio oppure il loro funzionamento non è adatto al contesto in cui devono operare. Lo sanno bene gli utenti che hanno tentato di usare il sistema di backup incluso in Windows con questi accessori: sulla carta entrambi funzionano a dovere ma nella realtà il Backup di Windows non riesce a scrivere correttamente sul disco di rete a causa dei problemi di compatibilità tra la modalità adottata dal programma di backup e quella prevista da Samba.

:: Sempre peggio

A questa situazione si somma l'ovvia conseguenza di avere un dispositivo teoricamente autonomo ma praticamente realizzato con un sistema operativo assolutamente privo di manutenzione. La mancanza di aggiornamenti, infatti, porta ad avere sulla rete un accessorio potenzialmente aperto a qualsiasi exploit e pur trattandosi di sistemi Linux, nativamente più sicuri di Windows, il pericolo non è certo da sottovalutare a cuor leggero. A questo si sommano, poi, problemi di compatibilità dovuti anche al modo in cui viene gestita la rete da Windows, perché anche questa famiglia di S.O. ci mette del suo: i protocolli di gestione della rete richiedono che lo username sui sistemi in comunicazione sia lo stesso. Il che significa che per avere una protezione valida ed evitare problemi con questi dischi di rete sia necessario replicare la login degli utenti dei computer anche sui dischi di rete, aumentando i tempi per il fatidico cambio di password. Già su sistemi casalinghi e SOHO, il cambio periodico delle password è rarissimo, figuriamoci se devono essere cambiate le password anche dei device di rete. In più, Samba non gestisce le password crittografate, costringendo a diminuire la sicurezza della rete Windows per poter realizzare una comunicazione efficiente. Nel complesso, l'introduzione di sistemi Linux based in una rete Windows SOHO è possibile ma, come tutte le cose, va fatta con criterio. Un criterio che i venditori non pubblicizzano, vantando questi accessori come semplicissimi da usare e senza necessità di manutenzione. Due punti di forza che qualunque esperto non può sottovalutare.

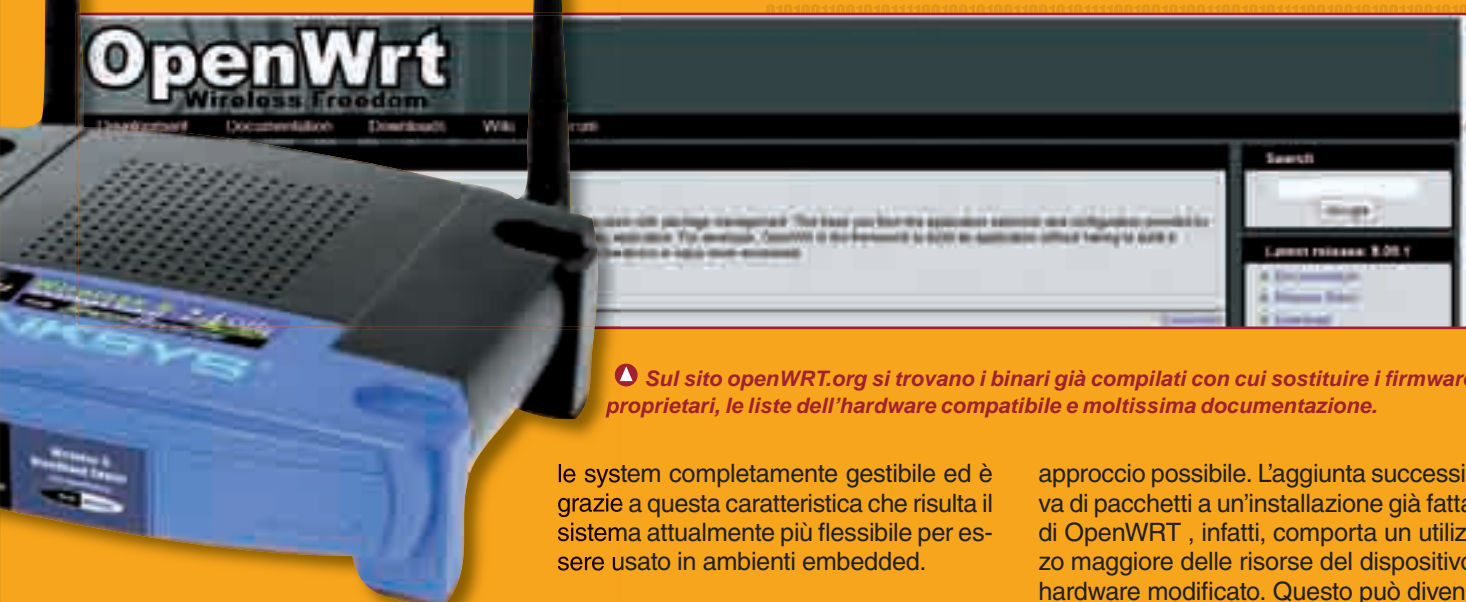
Un router più aperto

Non si sta parlando di infrangere le leggi: con OpenWRT possiamo dare un cuore Linux a tutti i dispositivi

A livello aziendale, i dispositivi di rete classici hanno subito un'evoluzione che li sta portando verso la loro trasformazione da device ad appliance. Basandosi tutti su tecnologie simili, è abbastanza logico che i router possano integrare man mano le funzioni dei firewall, quelle dei sistemi anti spam, quelle dei gateway e via dicendo. Questa evoluzione ha già dato buoni frutti in ambito Linux, con la creazione di alcune distribuzioni che permettono di trasformare qualsiasi computer in una appliance completa. Si tratta di soluzioni il cui hardware risulta costoso e che trovano poco spazio in applicazioni SOHO. A coprire questa fascia di mercato ci hanno pensato i partecipanti al progetto OpenWRT, openwrt.org: una distribuzione di Linux pensata esclusivamente per sistemi embedded.

:: Un ambiente

Più che un firmware dedicato ai device, OpenWRT si colloca a metà tra un ambiente operativo e un framework. Il suo utilizzo da parte dei produttori gli permette di basare ogni appliance su un ambiente standard che va a sostituire i classici ambienti proprietari, permettendo un abbattimento dei costi e un miglioramento generale sia delle prestazioni che dell'affidabilità. Lo sviluppo dei firmware, infatti, è la principale fonte di preoccupazione di qualsiasi produttore a causa dei problemi che ne possono derivare. Dal punto di vista degli utenti, invece, l'utilizzo di OpenWRT permette una personalizzazione assoluta e semplificata dei dispositivi. Nato per supportare soltanto i Linksys WRT54G, OpenWRT è stato ben presto espanso per fornire supporto a un'ampia gamma di dispositivi: dai telefoni Openmoko fino ai router Asus WL-500g. Attualmente supporta in modo nativo alcune funzioni tipiche dei router residenziali come il servizio DHCP e vari tipi di cifratura wireless ma si sta diffondendo sempre più per le sue caratteristiche avanzate: port forwarding, firewall, la possibilità di configurarsi come Wireless repeater, Access point o bridge, anche in combinazione tra loro, la possibilità di gestire un servizio di DNS dinamico on board, la gestione delle porte USB per l'aggiunta di dischi o la stampa di rete, la possibilità di monitora-

Sul sito openWRT.org si trovano i binari già compilati con cui sostituire i firmware proprietari, le liste dell'hardware compatibile e moltissima documentazione.

La serie WRT54G della Linksys è tra i dispositivi che ha avuto maggior successo in assoluto grazie alla sua flessibilità.

re in real time nello stato della rete e molto altro. In più, vera chiave di volta di tutto il sistema, il supporto all'installazione automatica di pacchetti ne permettono una ulteriore personalizzazione, rendendolo un vero e proprio sistema operativo. Diversamente dalla maggior parte dei sistemi proprietari, infatti, OpenWRT ha un file system completamente gestibile ed è grazie a questa caratteristica che risulta il sistema attualmente più flessibile per essere usato in ambienti embedded.

L'aggiunta di un disco esterno a un router con OpenWRT a bordo ci permette di avere un disco di rete sulla LAN.

:: Pronti, via!

Se si dispone di uno dei sistemi supportati da OpenWRT, il cui leneco è disponibile sul sito del progetto, le operazioni di installazione sono quasi banali. Abbiamo a disposizione sostanzialmente due tipi di file per l'aggiornamento: uno .TRX che contiene esattamente il firmware da inserire nel dispositivo oppure alcuni file .BIN che non sono altro che il primo file con l'aggiunta di un header. Questi file, disponibili per diversi modelli di hardware, permettono di seguire le procedure originarie di upgrade previste dai costruttori per sostituire il firmware proprietario con OpenWRT. Per comodità, anche considerando che il sistema non è obbligatoriamente rivolto ad esperti, conviene seguire le procedure dei produttori, così da garantirsi una buona riuscita dell'operazione. In caso di difficoltà, comunque, il sito del progetto spiega come procedere con aggiornamenti forzati in altri modi. Alla fine, comunque, un riavvio del dispositivo darà vita al nuovo firmware. Diversamente dalle versioni precedenti, nate strettamente per gli addetti ai lavori, le ultime versioni dispongono di una interfaccia web chiamata LUCI, che permette di configurare ogni aspetto standard del dispositivo. Naturalmente, se avremo aggiunto pacchetti al core di OpenWRT, LUCI non permetterà di configurarli e dovremo procedere tramite linea di comando, usando Telnet oppure SSH. Se fossimo esperti nella modifica di ambienti Linux, tuttavia, esiste un altro approccio possibile. L'aggiunta successiva di pacchetti a un'installazione già fatta di OpenWRT , infatti, comporta un utilizzo maggiore delle risorse del dispositivo hardware modificato. Questo può diventare un problema perché stiamo parlando di accessori che hanno sempre caratteristiche piuttosto limitate. Un'alternativa caldamente consigliata è quella di fare una injection di quello che ci serve all'interno del pacchetto di installazione, così da trasferire nel firmware del dispositivo il maggior numero di funzioni, preservando l'uso della RAM.

:: Limiti?

Avendo a disposizione un vero sistema operativo si potrebbe essere indotti a pensare di poter far svolgere qualsiasi compito a un dispositivo dotato di OpenWRT. Purtroppo non è così: le limitazioni dell'hardware sono veramnete forti e la vocazione del sistema è quella di consumarne il meno possibile, privilegiando il controllo delle connessioni. È quindi possibile usare OpenWRT su una scheda embedded che controlla un disco di rete ma non è questo il suo compito primario: altre distro potrebbero svolgere un compito migliore. Per ottenere buoni risultati sarebbe necessario ripulire il sistema da tutti i moduli di comunicazione che lo rendono così speciale. Malgrado queste considerazioni, comunque, l'hardware dei device più recenti è di tutto rispetto per i loro compiti e OpenWRT può facilmente arrivare a risolvere alcuni problemi, specialmente nell'applicazione ad ambiti particolari. In rete, sul Wiki e sul forum, sono disponibili decine di progetti già pronti che coprono ogni possibile esigenza di comunicazione.

ULTIMO BALUARDO

Normali impiegati, spesso annoiati, combattuti tra lavoro e famiglia: l'NSA, da dentro, non sembra eccitante

Nei giorni scorsi sono stato contattato da un vecchio conoscente che negli USA sta facendo carriera e non potevo rinunciare alla ghiotta opportunità: è un consulente dell'NSA, di origini italiane, che ha promesso alla moglie un viaggio nel tanto ambito Bel Paese. Oggi è giovedì, sono seduto a un bar nei pressi di Malpensa e la giornata non sembra tanto buona: è nuvoloso ma, soprattutto, la radio ha appena annunciato di un attentato contro la spedizione italiana in Afghanistan. Frank, lo chiamerò così anche se non è il suo vero nome, arriva in taxi, accompagnato da una moglie sorridente e da due bimbi piccoli piuttosto chiassosi. Non mi saluta nemmeno, a dir la verità: è al telefono e sempre parlando si siede poco più in là, confabulando in continuazione. Ne approfitto per dare il benvenuto in Italia al resto della famiglia. Un benvenuto lungo: passano due giri di colazione prima che Frank mi saluti decentemente. Con i bambini che hanno iniziato a giocare in un pantano lì accanto, la moglie a corrergli dietro e il telefono di Frank che squilla ogni 5 minuti, inizia una intervista che appare decisamente complicata.

Frank: Però, almeno, ci si vede. No?

HJ: Si, anche se speravo che la tua definizione di vacanza assomigliasse più al riposo che a una giornata infernale in ufficio.

Frank: Incognite del mestiere. Stamattina c'è stato un altro attentato e anche noi dobbiamo fare la nostra parte. Nel mio caso, poi, la situazione è particolare perché l'attentato riguarda anche il lavoro del mio ufficio.

HJ: A questo proposito: di cosa ti occupi esattamente? So che lavori per i servizi interni degli USA, che sei un ottimo informatico specializzato nei grandi sistemi ma non so nient'altro di quello che fai.

Frank: Purtroppo non posso essere più preciso e, forse, sai fin troppo. Non ti basta che sia uno simpatico e competente con cui chiacchierare? (Sorride)

HJ: Speravo in qualcosa di meglio. Comunque sia: mi spieghi un po' cosa fa l'NSA? Da quello che si dice in giro, dai film e anche dal tuo arrivo di stamattina, sembra che il lavoro lì dentro sia proprio qualcosa di eccitante.

Frank: A vedere certi film e a leggere

certe cose sembra proprio che io debba essere una specie di Rambo pronto a tutto, pieno di aggeggi elettronici… Ma non è proprio così. La National Security Agency è un'organizzazione che si occupa di tutelare la sicurezza degli USA sul territorio nazionale. Anche se il suo ambito operativo è quasi sempre interno, le attuali prospettive internazionali hanno un allargato il suo raggio d'azione anche all'estero.

HJ: Quindi viaggi all'estero, infiltrazioni, raccolta di notizie… Irruzioni nei sistemi informativi altrui?

Frank: Diciamo che può essere… Ma gran parte del lavoro è svolto semplicemente in ufficio. Da molti punti di vista, non c'è una grande differenza tra il mio lavoro e quello di qualsiasi altro impiegato di alto livello.

HJ: Immagino, però, che le cose di cui tratti non siano proprio banali…

Frank: Hai ragione. A pensarci, ogni giorno, ho a che fare con faccende che potrebbero essere molto costose, in termini economici o umani. Dall'NSA dipendono la sicurezza delle comunicazioni nazionali, la protezione delle personalità di spicco, l'attenzione verso insidie interne ed esterne. Gran parte del lavoro è dedicato alla prevenzione dei reati, specialmente quando si tratta di terrorismo. Il mio compito preciso all'interno della struttura non posso spiegartelo e, comunque, men che meno te lo farei scrivere. Quello che posso suggerirti è che l'NSA è uno degli organismi che nel mondo ha il più alto consumo di tecnologia. E di corrente elettrica!

HJ: Quindi niente modalità di risparmio energetico per i vostri computer?

Il memoriale dei crittologi, presso la NSA: nell'ombra hanno vinto più di una guerra.

Frank: (ride) Si, quella c'è... Anzi: siamo molto attenti alla questione ambientale e facciamo persino la raccolta differenziata. Però, come sai, dipende sempre da cosa accendi e i nostri server sono… Come dire… Ben messi.

HJ: A proposito di server ben messi… Echelon lo gestite voi, vero? Com'è? Puoi dirmi qualcosa di come funziona?

Frank: Echelon… Una cosa molto interessante per me ma non credo sia lo stesso per te o per i tuoi lettori. (Sorride)

HJ: OK, non ne vuoi parlare… Mi dici, almeno, se sai rintracciare le telefonate? Sai, sono un appassionato di NCIS…

Frank: (ride) Si, anche l'NSA lo fa. Prima che tu me lo chieda: non è il mio lavoro e nella finzione rendono complicate tante cose. Non ci vuole una vita per rintracciare una telefonata da fisso: giusto pochi istanti. Non serve nemmeno che la telefonata prosegua per preziosi minuti: basta sollevare il ricevitore per avvisare la centrale telefonica. Per le chiamate già terminate ci sono i registri degli operatori… Insomma: basta proprio poco. Ci vuole un po' più di tempo per le chiamate dai cellulari. Non posso spiegarti molto, però. Come ho già detto, non è il mio terreno.

HJ: Scendiamo sul tuo terreno, invece. Per quanto riguarda Internet?

Frank: Internet è un problema che viene tenuto sotto controllo. Diversamente dalle comunicazioni telefoniche, dove c'è sempre una centralina "istituzionale", in Internet puoi aspettarti un mucchio di sorprese. Hai un server di destinazione, un client da qualche parte del mondo e in mezzo può esserci di tutto: apparecchi di rete, tunnel, comunicazioni cifrate, proxy e un sacco di altra roba. Spesso iniziamo a lavorare su frammenti di comunicazioni di cui ignoriamo tutto e dobbiamo cercare di ricomporre un puzzle. Un problema piuttosto serio, dove il ragionamento è lo strumento per eccellenza. Se pensi che avere a disposizione computer potentissimi e una quantità di informazioni straordinaria possa essere d'aiuto, stai fresco. Il nostro è un lavoro duro, spesso frustrante. A volte ti puoi arenare per giorni, altre volte fila tutto liscio per scoprire di non essere approdati a nulla. Senza contare che, fuori

Una rara vista aerea della zona di Fort Meade in cui ha sede l'NSA.

dagli USA, le cose diventano politiche e sicuramente non puoi pretendere troppo.

HJ: Quindi partecipi anche ad azioni sul campo di qualche genere?

Frank: Io no. Il mio lavoro è strettamente di consulenza e sono stanziale. Ho una moglie e dei figli, una vita sociale…

HJ: Sei un consulente di un'agenzia misteriosa… Come si fa a coniugare la consulenza con la segretezza? Insomma: ogni giorno abbiamo davanti agli occhi esempi di consulenza che finiscono in liti, sistemi craccati, sicurezza violata…

Frank: Purtroppo il termine "consulente" ha due valenze. Da una parte c'è quello che può fare il lavoro per te mentre lavora anche ad altre cose, il professionista, la persona che vale molto. Dall'altro lato c'è il consulente che è tale perché non trova un'azienda che lo assume, un aspetto oggi, purtroppo, predominante.

HJ: Tu sei nella prima categoria?

Frank: Per mia fortuna si. Come, credo, tutti i consulenti di un certo livello che lavorano nell'agenzia. Io non voglio fare carriera nell'NSA, non mi interessa. Io ho una mia carriera lavorativa che sta andando verso livelli sempre più alti ma non voglio lavorare per l'NSA per tutta la vita. Quella è la differenza: l'agenzia prende come consulenti le persone che le interessano ma che non hanno interesse ad entrarci a tempo pieno. Non fa distinzioni di stipendio.

Il telefono squilla per l'ennesima volta e Frank cede: è ora di andare. Però offre lui, visto che il numero di interruzioni della nostra chiacchierata non si contano più. Il tempo di chiamare un taxi e di salutarci.

La nemesi della Pay-TV

Un sistema già vecchio che usa sistemi di protezione inconsistenti e già craccati: Mediaset Premium si vede gratis

Poco tempo fa abbiamo affrontato l'argomento della TV digitale, in occasione della presentazione della nuova piattaforma digitale satellitare, puntualizzando sul fatto che a nostro avviso si tratta di un servizio sostanzialmente inutile o, se non altro, ridondante: perché ci propinano nuove tecnologie, che richiedono nuovi decoder e nuovi impianti, quando ancora non è stato ultimato il passaggio al digitale terrestre su tutto il territorio nazionale? Qualunque sia la nostra opinione sull'argomento, e qualunque sia l'uso che ne facciamo nel nostro privato, ciò che abbiamo presentato era semplicemente un dato di fatto. In questo articolo, invece, vogliamo alzare un po' il tiro: vogliamo parlare un po' del sistema di protezione dei programmi criptati, e spiegare il perché è completamente inutile. Prima di affrontare l'argomento, però, ricordiamo un concetto fondamentale: le informazioni che qui riportiamo si basano su articoli già apparsi su Internet e dei quali non ci assumiamo alcuna responsabilità. Ciò che viene spiegato è assolutamente illegale, non va quindi fatto nella maniera più assoluta. A quelli tra noi che conservano un po' di etica personale, basta sapere che sono cose che si possono fare e ci servono solamente per capire che basare una protezione su concetti deboli è estremamente inutile. A ogni lettore la propria scelta e le proprie responsabilità.

:: Programmi protetti

La trasmissione dei programmi della televisione digitale non è molto differente a uno streaming video come quello che troviamo comunemente sul Web. In sostanza, anziché inviare un segnale radio modulato dal segnale analogico audio e video, sfrutto la frequenza per inviare dati digitali, che vengono poi convertiti in analogico dal decoder e inoltrati al sistema di visualizzazione (lo schermo televisivo). Il fatto che vengano trasmessi dati digitali rende facilmente implementabile un sistema di cifratura degli stessi, in modo tale da permetterne la decifratura e quindi la visualizzazione solamente a chi dispone di determinati re-

quisiti (leggi: ha pagato l'abbonamento). La verifica di questi requisiti non avviene da parte dell'emittente, ma direttamente nel decoder: il nostro credito si trova infatti memorizzato in una Smart Card che viene letta dal decoder stesso. Già questo sistema mostra tutta la sua debolezza: in assenza di controllo remoto, veder nascere schede clonate e manipolate è sempre stata solo questione di tempo, e ne abbiamo avuto la prova anni fa con i primi sistemi satellitari a pagamento. Nel caso del digitale terrestre è un po' più difficile; tuttavia il sistema adottato è caratterizzato da altre debolezze che analizzeremo tra breve e che possono essere usate da un cracker con sufficiente conoscenza e adeguata strumentazione per fare leva e scardinarne la protezione.

Mediaset Premium basa la tecnologia di protezione dei programmi sul sistema Nagravision, proprietario di un'azienda omonima. Nagravision, o meglio una delle sue prime implementazioni, è stato il sistema usato per codificare anche le trasmissioni analogiche ai tempi delle prime televisioni a pagamento, ma sostanzialmente, anche se notevolmente migliorato, il suo principio di base rimane oggi il medesimo di allora. Per non entrare troppo in dettagli, si tratta di "mescolare" le linee del fotogramma trasmesso in base a un numero pseudo-casuale ricavato da un determinato seed, usato come chiave e anch'esso trasmesso insieme al programma televisivo. Il decoder legge la disponibilità di credito sulla Smart Card, eventualmente scalando quando dovuto per la visione del programma, e in caso di credito sufficiente rileva il seed dal segnale ricevuto, lo usa per generare il numero pseudo-casuale necessario per decodificare il programma e prosegue con la decodifica stessa. La chiave, quindi, è proprio trovare il numero usato come seed, che equivale un po' a tentare di individuare una password o a craccare una chiave RSA.

Una scheda PCI in grado di ricevere la televisione digitale. Con un lettore di Smart Card, possiamo vedere anche quella a pagamento.

Sul Web

Come avviene spesso quando un'intuizione di qualcuno offre i natali a strumenti, spesso illegali, che vanno a toccare le tasche di qualche industria (come per esempio Mediaset), questi stessi strumenti hanno vita molto breve. Appaiono su un sito, vi restano il tempo sufficiente perché qualche autorità la individui e spariscono nel giro di qualche tempo. Tuttavia, sappiamo bene che tentare di mettere il bavaglio al Web e a Internet stessa è impossibile: ecco che, se cerchiamo con dovuta accortezza le informazioni di cui abbiamo bisogno (per esempio CAM, che sta per Conditional Access Module ed è il lettore di Smart Card dei decoder e dei moderni televisori, oppure deCSA, che è l'algoritmo che serve per decifrare il segnale) prima o poi le troviamo.

:: Programmi sprotetti

Il problema dei tempi delle trasmissioni analogiche criptate era che la tecnologia per leggere le Smart Card attraverso il computer non era alla portata di tutti. Per questo motivo chi voleva vedere a sbafo i programmi di TelePiù o simili ha dovuto attendere che qualcuno craccasse il sistema e comprare poi dal mercato nero una Smart Card sprotetta o un decoder modificato. Oggi però viviamo in un'epoca in cui proliferano i dispositivi interni o esterni per PC che possono ricevere le trasmissioni della TV digitale e permettono di vedere anche i programmi a pagamento, posto che il PC possa accedere alla Smart Card con un apposito lettore. Facciamo due più due: si dispone di un PC e di tutta la sua potenza di calcolo, questo PC può ricevere le trasmissioni digitali protette attraverso un'apposita scheda o a una chiavetta USB, si può quindi scrivere un programma che con un semplice brute forcing riesca a craccare la chiave usata per la generazione del numero pseudo-casuale e quindi a decodificare le immagini. E non è teoria: è già una realtà e sul Web si trovano diversi programmi e sorgenti che compiono il miracolo. Ovviamente, come abbiamo detto in apertura, si tratta di un'operazione illegale che qui riportiamo solamente a scopo didattico. La chiave usata per cifrare le trasmissioni digitali è composta da 48 bit, molto meno delle chiavi RSA usate oggi per proteggere altri tipi di comunicazioni su Internet, pertanto in capo a qualche tempo un cracker intenzionato a tutti i costi a sproteggere Mediaset Premium può essere certo che nel giro di qualche settimana avrà ottenuto il risultato sperato.

Un programma criptato visto senza decodifica ci appare come un insieme di linee "mischiate" e confuse.

Documenti protetti

Office permette di proteggere con password i documenti: che succede se perdiamo la perdiamo?

Quasi tutte le versioni di Office offrono la possibilità di proteggere i documenti creati usando una password e questa non è certamente parte di quel nutrito numero di funzioni inutili e sconosciute della nota suite da ufficio. A stare a vedere la quantità di servizi online per il recupero delle password di Office, infatti, sembra proprio che questo sistema di protezione abbia dato vita a un fiorente business. Il motivo, probabilmente, è che questo tipo di protezione risulta spesso troppo debole rispetto ad altri ma è anche quello più a portata di mano di qualsiasi utente, inclusi quelli tipici degli uffici, che hanno poca dimestichezza con le implicazioni derivate dall'uso di password.

:: Un business!

Nel corso delle varie versioni di Office, Microsoft ha adottato sistemi di protezione differenti e sempre più sofisticati che, tuttavia, hanno avuto spesso limitazioni o hanno sofferto di buchi di protezione dovuti a errori software. Il recupero della password dagli archivi creati con Office 97, per esempio, è pressoché istantaneo perché viene fatto bypassando il sistema di protezione e leggendo la password inserita nell'archivio. In altri casi, invece, il recupero di una password deve usare tecniche avanzate: in Office 2007, i file sono cifrati in modo forte e il recupero risulta più simile alla forzatura riservata alle password dei sistemi operativi o dei siti Web protetti. A differenza di queste password, tuttavia, i file di Office non subiscono conseguenze dai tentativi di decifrazione. Questo ha dato vita a una serie di servizi online che si occupano del recupero di questo tipo di file. Decryptum, www.decryptum.com, è un servizio di questo tipo che, in modo anonimo ed automatico, permette il recupero dei file di Word e di Excel. Il costo è tutt'altro che a portata di tutte le tasche perché se è vero che il recupero di un solo file costa meno di 30 dollari, utile in caso di emergenza per gli smemorati improvvisati, un recupero per 50 file arriva a costare quasi 500 dollari. Viceversa, il tempo di recupero è decisamente più

Decryptum, www.decryptum.com, è un sistema online dedicato per il recupero di password: si invia il file tramite il sito e si attende l'elaborazione. Da circa 30 dollari per file.

basso di quello impiegato da computer casalinghi e con metodi meno sicuri: un vantaggio che ha reso Decryptum il sistema più usato dalle aziende che si occupano di crimini informatici. Se pensiamo che fare le cose in proprio possa far risparmiare denaro abbiamo probabilmente ragione: l'acquisto di un programma per recuperare le password di Word e di Excel costa tra i 30 e i 130 dollari, in funzione delle sue caratteristiche. Si tratta, comunque, di prezzi di base per una licenza di prodotti che non offrono velocità particolari: usando un computer basato su Pentium 4, il miglior programma in circolazione riesce a provare qualche decina di migliaia di password al secondo. Per il recupero di una password scelta in base ai criteri di protezione consigliati dalla legge e dagli IT manager, i tempi di recupero arrivano a 2 settimane mentre, per le password più complesse, il compito diventa di improbabile soluzione.

:: Qualche calcolo

L'approccio brute force al recupero delle password di Office, infatti, risente degli stessi problemi del recupero di normali password.

Ipotizzando un computer in grado di tentare 500.000 password al secondo, l'ipotesi peggiore di recupero di una password che possa includere qualsiasi carattere ASCII e lunga 5 caratteri vede l'utilizzo della macchina per circa 4 ore. Considerando che la maggior parte delle password sono composte da lettere maiuscole e minuscole e sono lunghe da 7 a 9 caratteri, i tempi per il recupero sul nostro ipotetico computer variano da 23 giorni a 178 anni! Certamente, come per il password cracking tradizionale, l'uso di dizionari può semplificare notevolmente le cose ma non è detto che porti a una soluzione: se la password non è una parola standard ma un acronimo inventato in qualche modo, si rischia di attendere comunque molto tempo senza riuscire a concludere nulla, ricadendo nel caso del brute force classico. A causa di queste tempistiche lunghe sono stati inventati sistemi di recupero delle password che lavorano in modo distribuito: ElcomSoft Distribuited Password Recovery sfrutta la potenza delle GPU Nvidia in aggiunta a quella dei processori e può funzionare sfruttando una modalità parallela su anche 10.000 postazioni contemporaneamente. Ovviamente, le prestazioni e l'abbattimento di tempi di ricerca delle password si pagano: la versione di base costa la bellezza di 600 dollari.

:: Gratis!

Accanto a programmi che richiedono investimenti notevoli, esistono anche soluzioni gratuite che, tuttavia, non sono sofisticate e richiedono tempi di recupero lunghi. Il caso più eclatante riguarda Free Word and Excel Password Recovery. Si scarica dal sito www.freewordexcelpassword.com e supporta sia un attacco per dizionario che brute force. Il problema è che le sue routine non sono ottimizzate come quelle dei prodotti commerciali e la ricerca di password di Word o di Excel lunghe 7 caratteri risulta già impegnativa. Il consiglio per il recupero delle password dei file di Office, ovviamente, è quello di non averne mai bisogno: risulta sempre essere un costo in termini di denaro oppure di tempo impiegato.

CI PROVO?

Prima di iniziare il recupero di una password affidandosi al brute force, pensiamo se sia veramente il caso di provarci.

Per farlo, possiamo usare una formula specifica: (C^L)/S/N, dove C è la lunghezza della password, L è il numero di caratteri coinvolti ipoteticamente, S è il numero di password controllate ogni secondo da un computer e N è il numero di computer coinvolti. Ipotizzando un solo computer che controlla 50.000 password al secondo, riuscire a trovare una password di 8 caratteri che possono essere lettere maiuscole, lettere minuscole e cifre, la formula sarà: (8^(26+26+10))/50.000/1, pari a un numero di secondi piuttosto elevato: 196 seguito da 52 zeri.

Si può inserire Google Maps in qualsiasi sito Web: basta usare le sue API

UNA MAPPA NEL SITO

Diciamocelo pure: non sono disponibili mappe più precise di quelle offerte da Google Maps. La nuova versione delle API che Google mette a disposizione, poi, ha velocità di caricamento pari a quelle del sito originale e ci offre la possibilità di inserirle ovunque le riteniamo necessarie, con modalità abbastanza semplici. Diversamente dalle API precedenti, inoltre, non dovremo più nemmeno iscriverci per ricevere una chiave d'uso: le API versione 3 funzionano senza alcun problema anche senza registrazione. L'occasione, quindi, è ghiotta e ci permette di eliminare dai nostri siti quelle orribili scansioni dagli atlanti o quei sistemi statici e scontati che mostrano la posizione di aziende e negozi: Google Maps è più semplice per tutti e risulta, ormai, uno standard affermato.

:: Prima mappa

Basta dare uno sguardo al codice di esempio per capire quanto possa essere banale inserire una mappa. Nello specifico, la riga che ci mette a disposizione tutta la potenza di Google Maps è la prima inclusione di un codice Javascript, direttamente dal sito maps.google.com. Il richiamo alla API avviene passando come parametro la chiave sensor: serve per indicare all'intera API se il dispositivo da cui viene chiamata la pagina dispone di un sistema di posizionamento autonomo. La API, infatti, non serve solo per le pagine Web ma può essere sfruttata, in modo simile, anche all'interno dei nostri programmi e l'attivazione delle funzioni sensor permettono l'interfacciamento con un eventuale sistema GPS, esattamente come avviene se scarichiamo e installiamo sul cellulare l'applicazione Java Google Maps. Il Javascript contenuto nel corpo della pagina, invece, è quello che costruisce veramente la mappa che desideriamo. Nella prima riga viene creata un oggetto "mappa", puntato sulle coordinate 40,9. Nelle righe seguenti viene creata una variabile che sia in grado di gestire le opzioni desiderate per la mappa in fase di creazione. Nello specifico viene indicato un livello di zoom iniziale, viene specificato che deve essere centrata sulle coordinate fornite in precedenza e che il tipo di mappa da mostrare è

[Strappo 1]

```
<html>
<head>
<meta name="viewport" content="initial-scale=1.0, user-scalable=no" />
<script type="text/javascript" src="http://maps.google.com/maps/api/js?sensor=false"></
script>
<script type="text/javascript">
 function initialize() {
  var latlng = new google.maps.LatLng(40, 9);
  var myOptions = {
   zoom: 8,
   center: latlng,
   mapTypeId: google.maps.MapTypeId.ROADMAP
  };
  var map = new google.maps.Map(document.getElementById("map_canvas"), myOptions);
 }
</script>
</head>
<body onload="initialize()">
 <div id="map_canvas" style="width:100%; height:100%"></div>
</body>
</html>
```

stradale. L'ultima riga dello script serve per generare correttamente la visualizzazione all'interno della pagina, applicando le opzioni scelte in precedenza alla mappa nello spazio mapcanvas, definito inseguito, nel corpo dell'HTML.

:: Eventi, note, altro...

Abbiamo detto che le mappe così aggiunte al sito non sono statiche o banali: sono le mappe di Google! Questo significa, per esempio, che basta proprio poco per aggiungere segnali alla mappa creata.

[Strappo 2]

```
var marker = new google.maps.Marker({
   position: latlng,
   map: map,
   title:"Welcome!"
 });
```

Osserviamo un esempio nel codice contenuto nello **Strappo 2**: basta aggiungerlo a quello nello **Strappo 1**, dopo la riga che crea la variabile map, per ottenere un segnale, un marcatore, nel centro esatto della mappa, definito dalla variabile latlng. Il nome della mappa su cui inserire il marcatore è naturalmente definito dalla variabile map mentre la scritta da far comparire al clic è indicata nella definizione del marcatore stessa ma può essere costruita ad hoc, tramite javascript. Questo, tuttavia, è solo un primo passo perché basta poco anche per aggiungere funzioni avanzate. Immaginiamo di avere una stringa formattata HTML e chiamata contentString. Aggiungendo al codice prodotto finora quello contenuto nello **Strappo 3**, sempre dopo l'ultima riga del Javascript, potremo far apparire la stringa al clic sul marcatore definito in precedenza. La cosa inizia a farsi interessante: non solo possiamo inserire una mappa ma possiamo anche arricchirla facilmente con punti attivi di nostro interesse. Tuttavia non ci si ferma certo a questo: possiamo definire anche icone di genere diverso dallo standard, aree cliccabili, inserire note visibili in pagina e altro ancora. Con mappe di un certo livello di complessità, tuttavia, si rischia di esagerare con gli elementi ma non è un problema se organizziamo adeguatamente il lavoro: esattamente come le mappe originali, anche le nostre mappe personalizzate possono disporre di livelli, tramite una classe OverlayView. Richiamabile da pulsanti di controllo che possiamo inserire accanto alla mappa, questa funzione ci permette di mettere ordine nel materiale che disponiamo sulla mappa, arrivando a risultati eccezionali con sforzi minimi.

[Strappo 3]

```
 var infowindow = new google.maps.In-
foWindow({
   content: contentString
});

google.maps.event.addListener(marker,
'click', function() {
  infowindow.open(map,marker);
});
```

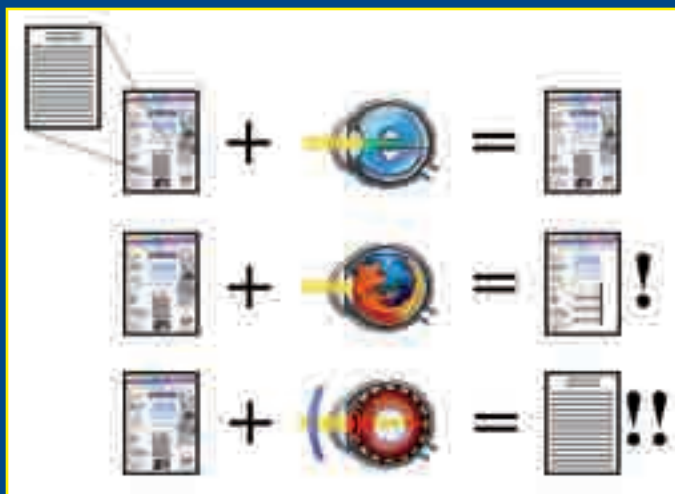
A proposito di interfaccia occorre dire che possiamo intervenire anche sui comandi disponibili all'interno della visualizzazione: settando a true l'opzione disableDefaultUI, i comandi di base delle mappe saranno invisibili. Potranno essere ottenuti se specificheremo come true, sempre tra le opzioni, le voci navigationControl, mapTypeControl o scaleControl: tutte e tre booleane e riservate, rispettivamente, ai controlli di navigazione nella mappa, ai controlli per il cambio del tipo di mappa e allo zoom. Le possibilità non sono ancora finite: le API di Google Maps includono anche sistemi di geocoding e di reverse geocoding. Da un indirizzo è possibile ricavare automaticamente le coordinate e viceversa. In questo modo risulta ancora più semplice costruire mappe personalizzate e in grado di adattarsi a qualsiasi situazione.

Web Scraping

"Rubiamo" informazioni ai ricchi per darle a tutti, in qualsiasi formato desideriamo

Information wants to be free è una frase diventata sempre più popolare, da quando Stewart Brand la pronunciò alla prima Hackers' Conference nel 1984. Con l'evoluzione del Web, infatti, gli utenti non sono più esclusivamente fruitori, ma anche autori di contenuti a disposizione di tutti: la conseguenza principale di questa evoluzione è che i dati hanno raggiunto un'importanza ben maggiore dei software che li generano, in quanto sono molto più difficili da replicare e richiedono la partecipazione di enormi comunità di persone. L'informazione, tuttavia, è tutt'altro che free, nel senso di gratuita: essa infatti vale fior di soldi per chi la raccoglie, ed è questo il motivo principale per cui la quasi totalità dei nuovi servizi "sociali" ci viene messa a disposizione gratuitamente. Allo stesso modo, essa è tutt'altro che free, nel senso di libera: infatti, basta provare ad accedere ad alcuni siti Web con browser che non siano i due o tre che vengono attualmente considerati standard per rendercene facilmente conto.
Come possiamo, quindi, liberare l'informazione? Nel primo caso, possiamo controllare i termini di servizio degli strumenti che usiamo, verificando che i dati che condividiamo vengano pubblicati con delle licenze aperte. Nel secondo caso, invece, si può fare molto di più: se dei dati raggiungono il nostro computer, infatti, noi possiamo filtrarli, elaborarli e alla fine visualizzare ciò che più ci interessa nel formato che preferiamo.

:: Familiarizziamo con la terminologia

E' proprio qui che entrano in gioco crawler e scraper: i primi (chiamati anche spider) sono programmi che navigano automaticamente all'interno di uno o più siti Web seguendo i link contenuti in ogni pagina che scaricano. I dati possono essere salvati integralmente oppure vengono filtrati dagli scraper, degli strumenti in grado di estrarre dal contenuto di una pagina esclusivamente quei dati che risultano interessanti per noi. Naturalmente, mentre è possibile fare un crawler generico (ad esempio quelli dei motori di ricerca, in grado di indicizzare pra-

Gli scraper ci mostrano le informazioni come desideriamo.

La open data cloud nell'aggiornamento di Luglio del 2009.

ticamente qualsiasi sito), è molto più complicato rendere generale il funzionamento di uno scraper: le informazioni da estrarre, infatti, possono cambiare a seconda della tipologia di sito o in base all'uso che se ne vuole fare.

:: Applicazioni

Le applicazioni possibili per uno scraper sono molteplici: ad esempio, scaricare contenuti da un sito Web, creare una sintesi e riproporla su una piattaforma differente; aggregare dati provenienti da fonti diverse salvando il tutto in un unico formato, come avviene nei sistemi di mashup; estrarre dati e farli analizzare ad altre applicazioni (ad esempio analisi statistiche su testi); oppure pubblicare automaticamente informazioni su un blog, o compilare dei form con dati generati automaticamente. I Web bot danno il meglio di sé quando le operazioni da svolgere, pur essendo semplici, sono lunghe e ripetitive.

:: Radiografia di uno scraper

Se desideriamo creare uno scraper è necessario innanzitutto avere ben chiara la sua architettura. Le operazioni principali che uno di questi programmi deve eseguire sono quattro: l'importazione dei dati (a sua volta, se vogliamo, divisa in crawling e scraping); la memorizzazione delle informazioni in un formato a noi comodo, ad esempio una serie di file in una directory, oppure un database; la loro elaborazione, ad esempio tramite aggregazione, analisi o conversione; infine, l'esportazione dei risultati o la loro pubblicazione su un medium specifico. Ognuna di queste operazioni richiede conoscenze specifiche che andremo ad approfondire a seconda del tipo di scraper che desideriamo creare. Ad esempio, se vogliamo pubblicare periodicamente sul nostro blog una tag cloud generata a partire dai post di un forum, dovremo studiare sia tecniche di analisi statistica dei testi per l'estrazione dei termini salienti sia le API messe a disposizione dal blog per la pubblicazione automatica.

:: Suggerimenti tecnici

Se vogliamo sviluppare uno scraper in pochi minuti, ci basta cercare tutorial specifici per il nostro linguaggio di programmazione: ormai ce ne sono di ogni tipo e sicuramente ci renderanno operativi in brevissimo tempo. Tuttavia, le variabili da prendere in considerazione quando si crea uno scraper sono moltissime e nessun tutorial è in grado di prevedere tutti i problemi che ci si possono parare davanti. Per fortuna, ci sono alcuni concetti comuni che costituiscono un ottimo punto di partenza: questi concetti riguardano la tecnologia che accomuna ogni scraper, cioè il Web stesso. Per partire con il piede giusto, quindi, è necessario prima di tutto capire con cosa abbiamo a che fare: principalmente pagine Web e form (quindi codice HTML), collegamenti fra il nostro browser e un server (quindi protocollo HTTP), estrazioni di stringhe da un testo (quindi regular expression). Per HTML ed HTTP possiamo controllare le relative RFC, mentre per le regular expression troveremo un ottimo tutorial all'indirizzo http://perldoc.perl.org/perlretut.html. Un altro manuale, un po' datato ma più generale, è disponibile all'URL http://davide.eynard.it/malawiki/PowerBrowsing.

:: Conclusioni

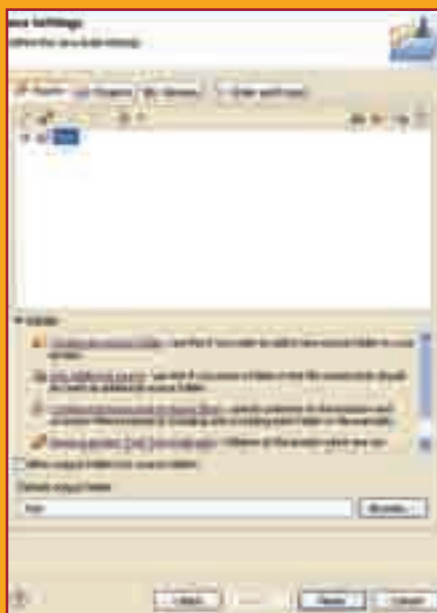
A questo punto, rimangono solo un paio di accorgimenti per chi di noi vuole cimentarsi nello sviluppo di scraper: il primo è quello di creare un bot che segua anch'esso le regole della netiquette (http://www.december.com/cmc/mag/1997/feb/helspid.html); il secondo è decidere come condividere le informazioni che abbiamo collezionato. Se i dati che abbiamo raccolto hanno una licenza che consente di condividerli liberamente, una delle soluzioni più interessanti potrebbe essere quella di condividerli come Linked Data (http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData/): in questo modo entreranno a far parte di una "nuvola" di dati già liberi e collegati fra loro, a disposizione di tutti per ogni tipo di utilizzo.

Un'ottima soluzione che aiuta a creare programmi per qualunque piattaforma

Dedicato ai programmatori

Non è la prima volta che HJ si occupa di ambienti di sviluppo dedicati ai programmatori di ogni livello, ma abbiamo sempre cercato di offrire alternative valide e gratuite ai pesanti dinosauri commerciali come Visual Studio, del quale abbiamo preso in considerazione solo la versione Express in quanto quelle professionali hanno costi che sono ben al di fuori della portata di molte persone che si occupano di programmazione solo per diletto. Eclipse è un'altra di queste valide alternative, l'unica probabilmente a essere veramente indipendente dalla piattaforma su cui è eseguita e da quella target, cioè quella per cui vogliamo sviluppare programmi. Non è un IDE come molti altri, ma un insieme di tecnologie completo, gratuito ed estremamente versatile che contiene anche strumenti IDE e che (perdonateci il gioco di parole, ma ne siamo certi) finirà per eclissare gli altri di cui abbiamo già parlato.

:: Che cos'è Eclipse

Prima di precipitarci sul sito e scaricare tutto quanto è possibile, dobbiamo stabilire che cosa vogliamo sviluppare e per quale piattaforma. Di Eclipse, infatti, esistono diversi pacchetti, ognuno dedicato a un particolare ambito di programmazione. Per esempio, possiamo scaricare il pacchetto PHP e creare meravigliose applicazioni Web, ma disporremo solo degli strumenti dedicati a questo ambiente. Eclipse in effetti è una raccolta di progetti basati su Equinox, ognuno dedicato a un ambiente specifico o alla soluzione di un problema ben determinato, ed è quindi difficile scegliere che cosa scaricare. Ricordiamoci che non si tratta di un comune IDE: al livello più basso, non avremo a disposizione un editor per il codice e strumenti di correzione e programmazione classici, quali quelli che siamo abituati a vedere in altri prodotti. È infatti l'insieme degli strumenti integrati in ogni distribuzione che rende Eclipse un pacchetto adatto per ogni particolare compito. L'architettura su cui si basa Eclipse è quella offerta da Java: per questo motivo tutto il pacchetto è altamente portabile e può funzionare su qualunque sistema. La sua architettura è basata su plugin, pertanto possiamo scaricare un pacchetto completo già predisposto per un particolare compito, oppure scaricare un pacchetto base e ar-

La creazione di un nuovo progetto è guidata: dopo aver scelto i percorsi per i file, ne possiamo definire la struttura.

ricchirlo a mano a mano con i plugin degli strumenti che ci possono tornare utili o che sono essenziali per il nostro lavoro e renderlo quindi un ambiente di sviluppo personalizzato e ritagliato intorno alle nostre esigenze. Unico requisito è la presenza delle librerie Java già installate sul computer che intendiamo usare: le troviamo come al solito sul sito Sun www.java.com/it per tutti i sistemi operativi.

:: Download e installazione

La home page di Eclipse è all'indirizzo www.eclipse.org; qui troviamo i collegamenti per informazioni, documentazione e download per tutto quello che riguarda il progetto. Se facciamo clic su Download, ci possiamo rendere conto facilmente di quanto sia versatile Eclipse e di quanti progetti e prodotti ruotino attorno a questa tecnologia. Se abbiamo già le idee chiare, possiamo scegliere il download che più si adatta alle esigenze del momento (per esempio Eclipse IDE for C/C++ Developers), altrimenti possiamo provare a scaricare il pacchetto classico che ci offre la possibilità di sviluppare per Java, e arricchirlo poi in base alle nostre necessità. In queste pagine parleremo principalmente del pacchetto classico (Eclipse Classic 3.5.0) e di quelli principali e più adatti a un utilizzo generico, ma i concetti che valgono per questi valgono poi per tutta l'architettura di Eclipse. Il download per Windows è costituito da un semplice file Zip che va estratto direttamente nella cartella di destinazione. Facciamo attenzione alla lunghezza del nome della cartella: dato che Windows supporta al massimo una lunghezza di 255 caratteri per i nomi dei file completi di percorso e che Eclipse contiene numerose sottocartelle annidate, se usiamo un percorso lungo si verificheranno facilmente errori. Meglio quindi estrarre il file in una cartella nella root del disco C oppure, se vogliamo le cose più ordinate, usare al massimo C:\Programmi\Eclipse. Possiamo poi creare sul desktop un collegamento che punti al file eclipse.exe che si trova nella cartella del software, per facilità di utilizzo. La prima cosa che dobbiamo fare all'avvio del programma è scegliere il workspace da usare: possiamo crearne diversi, uno per ogni progetto (in quanto progetti diversi potrebbero aver bisogno di moduli diversi), oppure lasciare come opzione predefinita quello proposto, nel caso in cui non svilupperemo sicuramente per altri ambienti.

All'avvio di Eclipse possiamo scegliere quale workspace usare: una comodità quando usiamo l'ambiente di sviluppo per più progetti diversi.

La schermata di avvio offre collegamenti all'ambiente di lavoro, alla documentazione e ai tutorial disponibili.

:: Utilizzo

A questo punto siamo già pronti per lavorare, ma dobbiamo tenere presente una cosa: dato che ci troviamo in un IDE, e soprattutto in un IDE di livello elevato, non possiamo trovarci "pulsanti magici" che creano l'applicazione per noi. Un nuovo progetto è sempre un nuovo progetto desolatamente in bianco, che contiene solamente gli elementi essenziali per il suo avvio e i suo funzionamento, ma tutto ciò che farà lo dovremo scrivere noi di sana pianta e con tanta pazienza. Dobbiamo quindi conoscere le tecniche di programmazione e la semantica del linguaggio che stiamo usando o non andremo da nessuna parte.

Prima di partire in quarta con l'uso di Eclipse, conviene spendere un po' di tempo a studiarne la documentazione. Essendo diverso da altri IDE con cui possiamo aver avuto a che fare, molti strumenti non sono dove ci aspettiamo di trovarli (o non sono proprio presenti) e soprattutto funzionano in modo diverso a quanto siamo abituati. Sul sito troveremo abbondanti informazioni, ma vale la pena fare un giro anche nella sezione Wiki (http://wiki.eclipse.org/Main_Page) dove possiamo apprendere molto sull'architettura e sull'utilizzo di questo meraviglioso ambiente di sviluppo.

Fra LED, biotecnologie e divulgazione scientifica, 30 anni di Ars Electronica

The "Human Nature"

Linz è una piccola cittadina austriaca. Apparentemente anonima e del tutto ordinaria, non solo Linz prospera, ma è diventata una meta obbligata per artisti, intellettuali, curatori, ricercatori e sperimentatori di tecnologie applicate all'arte da tutto il mondo.
Ad attirarli è Ars Electronica, il più grande festival a livello europeo dedicato all'arte digitale. E in effetti in molti anche dall'Italia si sono spostati viaggiando su treni e aerei fra il 3 e l'8 settembre per il trentennale della manifestazione, un'edizione in cui ben due progetti italiani hanno ricevuto una Honorary Mention nella sezione "Digital Communities": Wikiartpedia di Tommaso Tozzi e HackMeeting (congratulazioni da parte nostra agli hackari per aver suggellato il loro ingresso nell'olimpo dell'arte digitale).

:: Antropocene. Una nuova era

Ars Electronica si confronta quest'anno col tema "Human Nature". I suoi curatori sostengono che stiamo entrando in una nuova era caratterizzata dall'influenza irreversibile dell'uomo sull'ambiente: l'Antropocene. Ma non è nostra abitudine dare giudizi su un evento artistico a cui non abbiamo partecipato. Per questo abbiamo chiesto cosa ne pensava Simona Lodi, art director del Piemonte Share Festival, che al contrario ha seguito il festival dall'inizio alla fine.
Ecco il risultato della nostra chiacchierata virtuale. "A Linz si respirava un'aria diversa, che risente ancora di un momento di fiducia nello sviluppo tecnologico e nel benessere economico. La tematica Human Nature non è stata però tra le più brillanti degli ultimi anni. Riecheggiava quella di Hybrid, dando centrale importanza all'ingegneria genetica e alla biotecnologia, sfiorando appena le questioni etiche, strizzando l'occhio alla medicina avanzata e contrapponendo il classico dualismo tra Nature and Nurture, senza averne uno sviluppo ecosostenibile e comunque non critico. Dopo 30 anni la missione di Ars Electronica rimane la stessa: mettere in dialogo

l'arte con la tecnologia e la società in un futuro che usa sempre il metodo scientifico come lente per vedere l'essere umano e il mondo. Senza entrare nello specifico della manifestazione, ho trovato interessante il progetto 80+1, un progetto globale sulla collaborazione umana detto "cloud intellegence". Tra le curiosità dell'archivio storico delle celebrazioni 30ennali mi ha divertito vedere un'intervista di un giovane Derrick de Kerckhove e di un ancor più giovane Joi Ito che lo guardava con aria spaventata. Mentre mi ha colpito l'attaulità del lavoro Life Writer di Laurent Mignonneau & Christa Sommerer del 1996. E la mostra "see this sound" sul legame tra suono e immagine nell'arte, una collaboration fra il Lentos Art Museum Linz e il Ludwig Boltzmann Institute Media.Art.Research, molto interessante. Non mi sono bastate 3 ore per vederla tutta..."
Torneremo nelle conclusioni su alcuni punti delicati di questa analisi.

:: Il Centro

Ma Ars Electronica da quest'anno è anche Ars Electronica Center, un museo/laboratorio inaugurato da poco, frutto di un investimento economico imponente e di lunga durata: 3000 mq dedicati alle esposizioni; 400 per seminari e conferenze; 1000 a R&D (laboratori di ricerca e sviluppo); 650 a bar e catering service; infine una piazza di circa 1000 mq serve per organizzare eventi "open-air". È difficile immaginare progetti del genere in Italia, dove la ricerca non viene certo premiata e le arti digitali si trovano a competere con Donatello, Michelangelo e il Colosseo: una lotta impari. Simona, che ha avuto occasione di visitare il nuovo centro e che ha dalla sua parte una lunga esperienza di curatrice, sembra la persona adatta a cui passare un'altra volta la palla: la sua risposta è tagliente.
"In Italia i musei ci sono già. Per quello che riguarda il contenitore abbiamo tanti edifici storici meravigliosi che potrebbero funzionare anche meglio dell'Ars Electronica Center, che comunque è un edificio di 6500 mq ed è costato 30 milioni di euro e che ha solo una facciata a LED luminosi spettacolare. Diverso è il progetto museale, che comunque è un contenitore per tecnologia avanzata e non per l'arte in epoca digitale. Ospita anche mostre d'arte, ma non è dedicato solo a quello. Lo scopo principale è un museo divulgativo sulle tecnologie applicate alla biologia e alla medicina. Bello sarebbe invece in Italia allargare le collezioni di musei d'arte contemporanea anche alla newmedia art. Ma per ora non vedo nulla qui da noi." Felici di essere stati spiazzati, questa ipotesi del tutto percorribile ed "ecologica" in senso lato, ovvero basata sul riuso e sul potenziamento delle strutture museali esistenti e sulla creazione di zone ibride, finisce per annullare un falso dualismo, una concorrenza fra discipline classiche e sperimentazione digitale che potrebbero sostenersi a vicenda: anche a fronte di tagli e ristrettezze economiche, la parola chiave sembra come sempre apertura (nel senso più ampio del termine), accoppiata all'intelligenza nella gestione delle risorse.

Ars Electronica: etoyMAYA presenta il panel "The Human Nature".

Ars Electronica Center. Peccato che in Italia sia un'ipotesi utopistica.

:: Conclusioni

Una Ars Electronica che punta sulle biotecnologie, capace di attrarre capitali e investire in ricerca senz'altro, ma priva di uno slancio critico profondo, quella che ci racconta Simona Lodi e guardando dall'esterno ci sentiamo in linea con la sua analisi. Aggiungendo due note critiche di carattere generale. Da un lato, l'esistenza di correnti filosofiche e artistiche che portano all'estremo la trasformazione dell'uomo in senso biotecnologico – i transumanisti, ad esempio – rappresentano un terreno fertilissimo in cui la possibilità di riesumare teorie vecchie e mai sopite come il "superuomo" assume nel contemporaneo un realismo inquietante. Dall'altro, la tendenza a ricondurre l'intervento artistico/tecnologico alla spettacolarità significa in ultima analisi neutralizzare le potenzialità radicalmente evolutive (a livello politico, sociale e antropologico) delle tecnologie digitali, rappresentandone il lato più marcatamente propagandistico. E non solo.

penelope.di.pixel

Un Nokia pulito

Se abbiamo un cellulare personalizzato da un operatore, possiamo ripulirlo con un semplice trucchetto

La scena è delle più classiche: un cellulare si rompe e si decide di comprarne uno nuovo, magari appena uscito, recandosi in uno dei tantissimi negozi disponibili. Poi la sorpresa: il cellulare è costato quanto il listino comanda ma fin dalla sua accensione mostra il logo di un operatore, un menu personalizzato e dispone già di configurazioni di vario genere. Se l'operatore corrisponde al nostro, la cosa può anche essere gradevole ma se l'operatore è diverso sa di presa in giro ed è normale chiedersi come mai si è pagato decine o centinaia di euro per un cellulare che sembra taroccato in partenza. Un problema ancora più stringente quando è praticamente impossibile trovare in circolazione versioni sane di alcuni modelli. Naturalmente, il giochetto vede coinvolti gli operatori telefonici e i produttori di cellulari: questi ultimi ricevono certamente qualche sovvenzione, diciamo così, per produrre modelli ad hoc per gli operatori. Modelli che sono del tutto simili a quelli standard e si differenziano, appunto, per la presenza dei loghi di un determinato operatore, delle sue configurazioni, dei suoi programmi. Così si finisce per usare un cellulare TIM con un abbonamento Vodafone o viceversa, facendo molta, moltissima confusione. Come se non bastasse, il firmware degli operatori è spesso più limitato di quello standard, ammette meno personalizzazioni e viene aggiornato con una priorità inferiore. Un disastro le cui spese vengono fatte dagli acquirenti finali, sacrificati in nome della pubblicità. Si, perché non si sta parlando di cellulari comprati d'occasione che hanno la cosiddetta sim lock: si sta parlando di cellulari senza blocchi, pagati a prezzo di listino e regolarmente acquistati in negozio e utilizzati fin da subito con qualsiasi operatore.

:: Il trucco

Un aspetto spiacevole che si rischia di pagare con l'acquisto di qualsiasi cellulare di ogni marca

Sul sito hunajatehdas.net/nokia/firmware/ en sono disponibili i product code di qualsiasi telefono cellulare Nokia.

Stesso hardware ma prezzi diversi: basta qualche minuto per trasformare un nomale nokia N73 in un N 73 Music Edition.

ma che risulta di facile soluzione se si è acquistato un Nokia. I numeri dei modelli di cellulari di questa azienda, infatti, sono composti da cifre che segnalano esattamente cosa abbiamo tra le mani grazie a una codifica parlante. In questo modo, a ogni codice corrisponde un modello, un colore, un sotto modello, un operatore per cui è stato personalizzato e via dicendo. Per esempio, un Nokia N80 personalizzato per TIM ha come product code 0537357 mentre lo stesso identico cellulare che non ha alcuna personalizzazione ha product code 0529381 se la cover è nera o product code 0529378 nel caso la cover sia argento. Per rimuovere le personalizzazioni, quindi, l'aggiornamento standard del firmware non serve a nulla: il sistema di aggiornamento Nokia invia al server il product code del telefono per recuperare il firmware corretto che, ovviamente, sarà sempre personalizzato dall'operatore. Togliersi di torno questa personalizzazione è questione di un attimo: bisogna cambiare il product code del telefono, inserendo quello standard oppure, se vogliamo, quello che Nokia ha realizzato per il nostro operatore. In questo modo, il telefono sarà riconosciuto come standard oppure con il brand "corretto" e l'aggiornamento automatico del firmware eliminerà le personalizzazioni indesiderate. Il primo passo da fare è quello di installare gli strumenti Nokia forniti col telefono per fare un backup completo: nella maggior parte dei casi, questa operazione elimina totalmente i dati. Poi dobbiamo installare un programma chiamato Nemesis Service Suite, scaricabile gratuitamente dal sito www.b-phreaks.co.uk. Durante l'installazione del programma, scegliamo di usare il driver virtuale USB, in modo da fargli rilevare il nostro telefono cellulare. Poi dobbiamo avviare Nemesis, cliccare su Scan for New Devices, selezionare Phone Info, cliccare sul pulsante Scan e poi su Read. Ora compariranno nella finestra di destra alcuni numeri di riferimento del telefono. Annotiamo il Product Code e prepariamoci a cambiarlo. Apriamo un browser, visitiamo il sito hunajatehdas.net/nokia/firmware/en e cerchiamo nella casella a sinistra il modello del nostro cellulare. In basso alla finestra verrà riportato il product code della versione no brand. Scriviamolo in Nemesis al posto di quello che ci era stato segnalato, spuntiamo la casella Enable accanto al numero appena inserito e clicchiamo su Write. A questo punto Nemesis non ci servirà più, così come potremo chiudere il browser: il telefono risulta a tutti gli effetti un telefono no brand. Ora possiamo procedere con l'aggiornamento standard del firmware, tramite Nokia Data Suite. Se dovesse venire rilevato un firmware della stessa versione di quello installato, scegliamo di sovrascriverlo: la versione sarà pure la stessa ma la personalizzazione dell'operatore verrà fatta sparire.

:: Fare o evitare?

Il trucco è di facile realizzazione e ci vuole più tempo per spiegarlo che a farlo. In più ha un vantaggio enorme: in caso di errore di scrittura del numero, è possibile ricorrere nuovamente a Nemesis per riprovare la modifica. Certamente c'è la possibilitào di bloccare il telefono in via definitiva ma è un rischio che va comunque valutato quando si ha a che fare con aggiornamenti del firmware, anche se fatti in modo istituzionale. In più, l'operazione invalida la garanzia. Prima di chiedere una riparazione sarà probabilmente meglio ripristinare il software dell'operatore intruso. D'altra parte, questo tipo di hack è molto in voga e dà vita a risultati piacevolmente inattesi, specialmente quando si ha a che fare con modelli identici che differiscono solo per i programmi disponibili pur avendo prezzi di vendita differenti. Il caso del Nokia N73, per esempio, è tra i più evidenti: seguendo la procedura di eliminazione del brand di un normale Nokia N73 ma inserendo il product code di un Nokia N73 Music Edition, dopo 'aggiornamento del firmware ci si ritrova con quest'ultimo modello di telefono invece del modello classico.

Nemesis Service Suite è un software per la lettura e scrittura di memorie a basso livello: indispensabile per la modifica di qualsiasi telefono BB5 come, appunto, i Nokia.